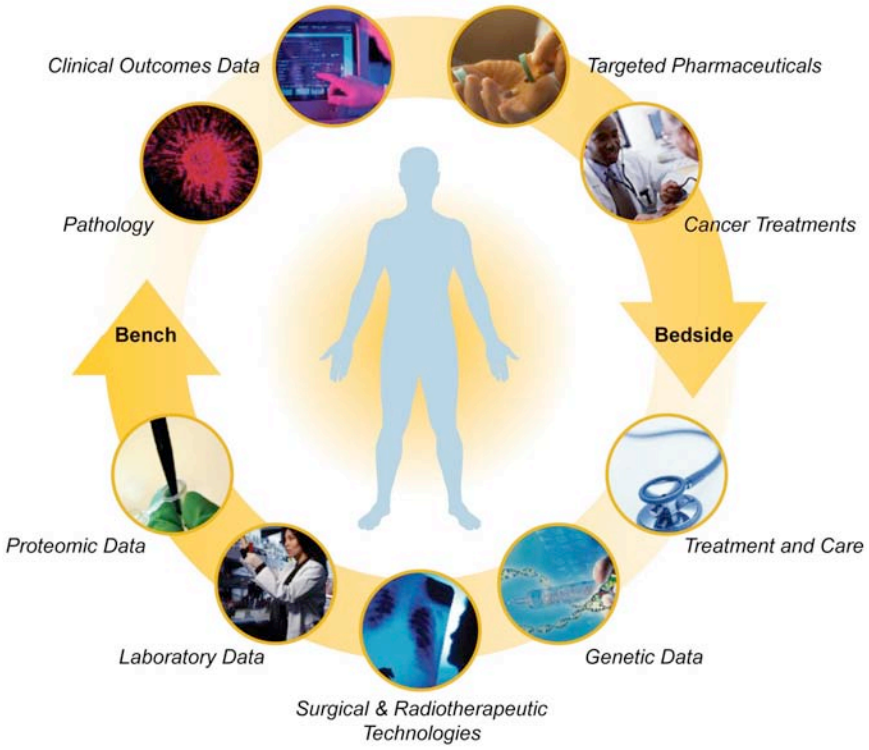
The Role of caBIG® in the New Era of Biomedicine
Kenneth H. Buetow, Ph.D.
September 13, 2010

The Goal: Medicine That is Personalized, Predictive, Preemptive, Participatory

• Unifies discovery, clinical research, and clinical care (bench-bedside-bench) into a seamless continuum
• Results in improved clinical outcomes
• Accelerates the time from discovery to patient benefit
• Embraces the global nature of disease and the unique international insights in addressing its challenge
• Utilizes Health 2.0 technologies
• Engages all stakeholders
• Enables a Rapid Learning Health System



What Slows the Transformation

- Tsunami of Genomic and Clinical Data
- Islands of Information
- Standard Language
- IT Systems Do Not Interoperate

caBIG® Overcomes the Obstacles to Data Integration

- 70+ Software Tools to overcome Tsunami of Genomic and Clinical Data
- Standardized Vocabularies to overcome Standard Language
- National Grid for Data Sharing to overcome Islands of Information
- Interoperability to overcome IT Systems Do Not Interoperate

caBIG® Initiative: Core Components

- Community
  - All sectors of the community are welcomed to participate in the concepts, development, and usage of capabilities.
- Connectivity
  - All stakeholders are to be connected with an infrastructure that permits rapid and seamless transfer of information.
- Content
  - Users are to be provided with the data they need for research and clinical care, in order to accelerate discovery and improve patient outcomes.

caBIG® At-A-Glance

- **Community**
  - 2,300+ participants
  - 700+ organizations
  - 15+ countries
  - 19 licensed Support Service Providers
  - 1,100+ attendees registered for 2010 caBIG® Annual Meeting
- **Connectivity**
  - 78+ applications
  - 149 “nodes” connected to National Grid via caGrid
- **Content**
  - 2.17 million biospecimens available through caGrid
  - 4.76 million images stored in National Biomedical Imaging Archive
  - 39,952 microarray experiments available on caGrid

Institutional Affiliations of Individuals Registering to use the REMBRANDT GBM molecular analysis portal since September 2010

Columbia University, USA
Inserm, FR
University of Bergen, NO
Aichi Cancer Research Institute, JP
Dana-Farber Cancer Institute , USA
Johns Hopkins School of Medicine , USA
University College London , UK
National Taiwan University , TW
Eli Lilly & Co
Clinic University of Giessen , CN
DKFZ , DE
University of Pittsburgh, USA
University of Houston, USA
Henry Ford Health Systems , USA
MD Anderson Cancer Center , USA
Brandeis University , USA
Duke University, USA
Stanford University School of Medicine , USA
North Carolina State University, USA
University of Calgary , USA
HHMI, University of Pennsylvania, USA
BIDMC, Harvard University, USA
Cal. Pacific Med. Center Research Inst. , USA
University of Antwerp, BE
CPMC Research Institute, USA

caBIG® In Action: Powering a New Generation of Research

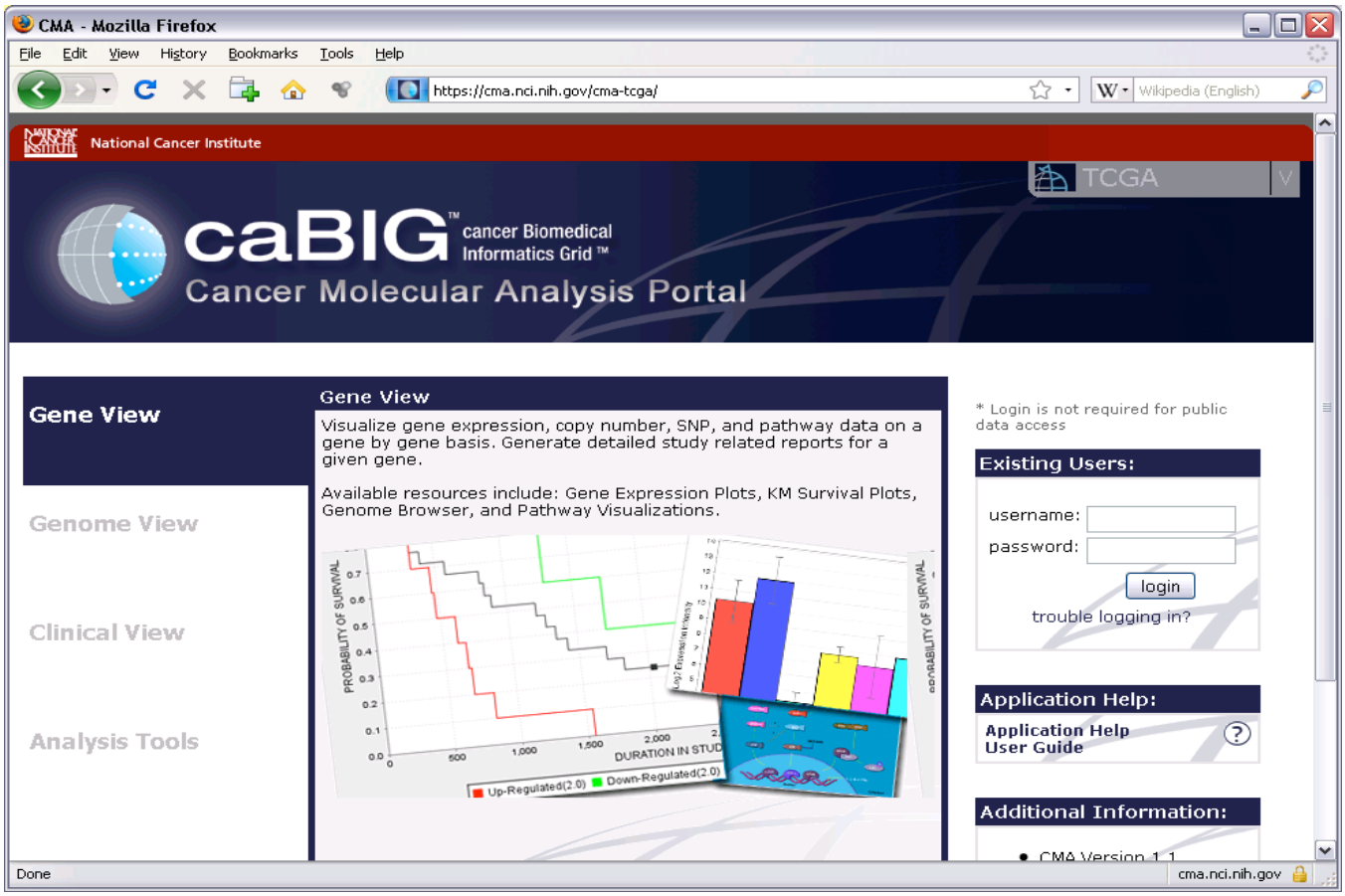
caBIG®: integrating multidimensional molecular data

caBIG® empowers researchers to see the “BIG” picture by integrating increasingly complex layers of cancer biology, from gene to clinical phenotype, as a whole:

From Gene to Genome to Genomes to Pathways to Clinical Outcomes

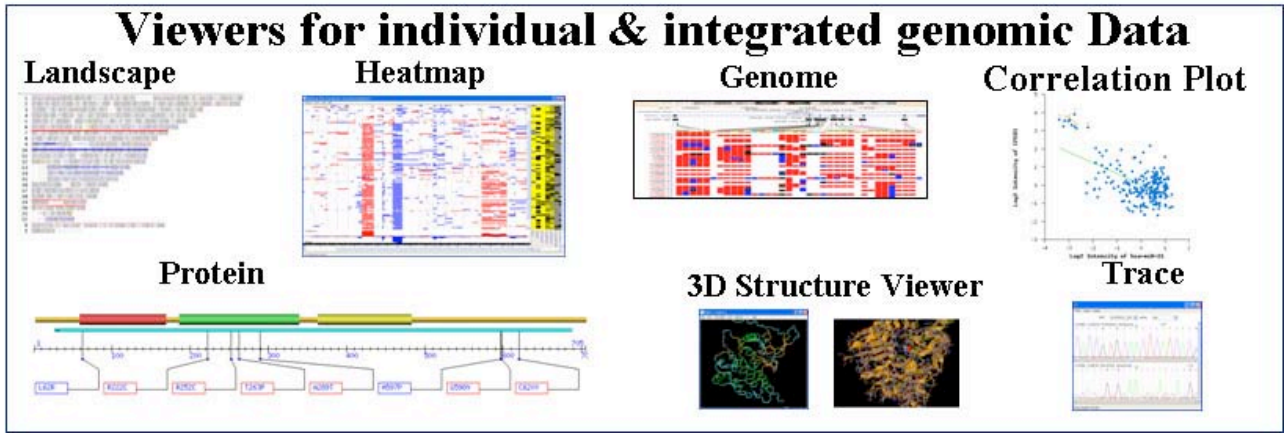
All from their computer

A an in silico cancer research portal (http://cma.nci.nih.gov)



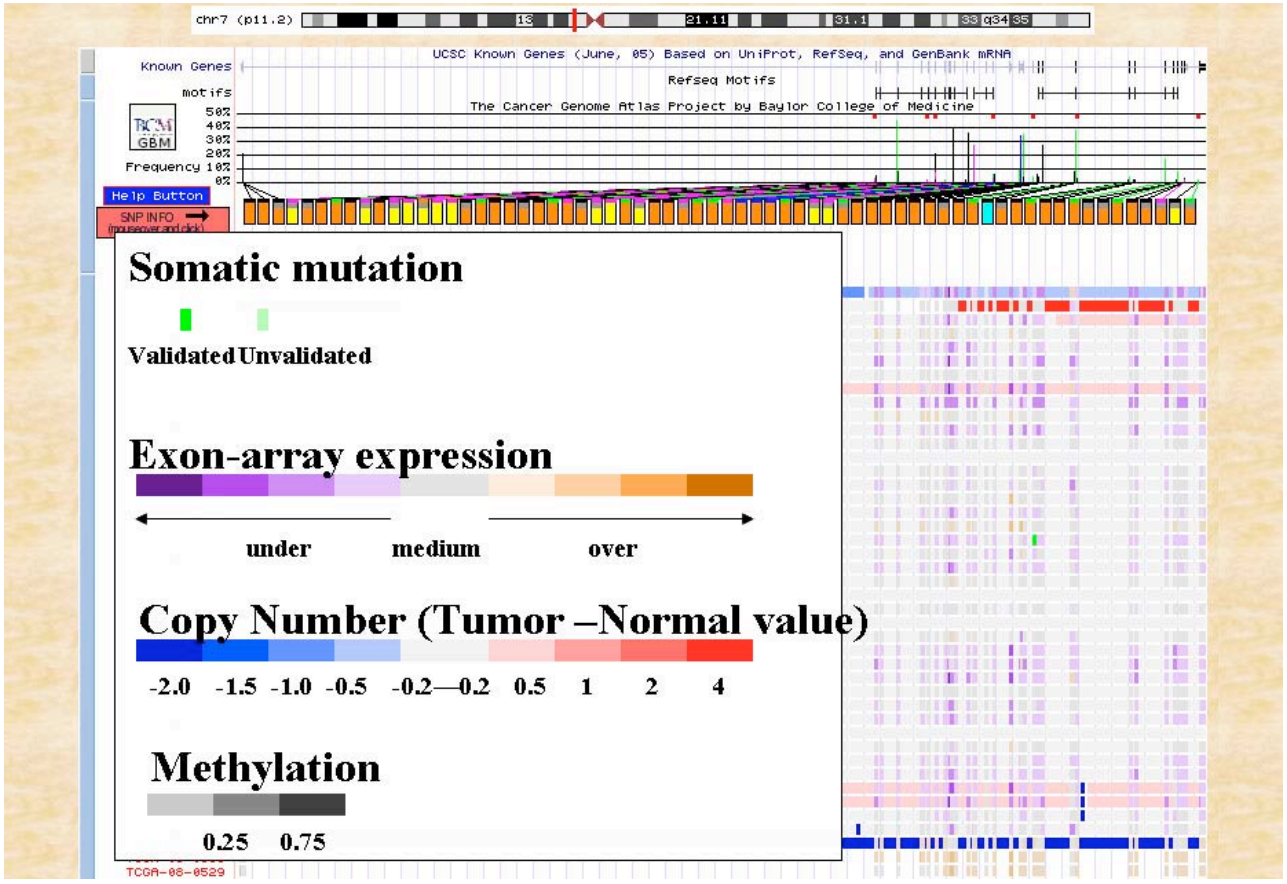
Cancer Genome Workbench (http://cgwb.nci.nih.gov)

Project	Disease	Genomic Data				
		Copy Number	Gene Expression	Methylation	Clinical	Mutation
TCGA	GBM, Ovarian	+	+	+	+	+
TSP	lung	+			+	+
TARGET	ALL, NB	+	+		+	+
JHU	GBM, Pan					+
Rembrandt	GBM	+				
GSK cell line	>30 tissues	+	+			
COSMIC	>30 tissues					+

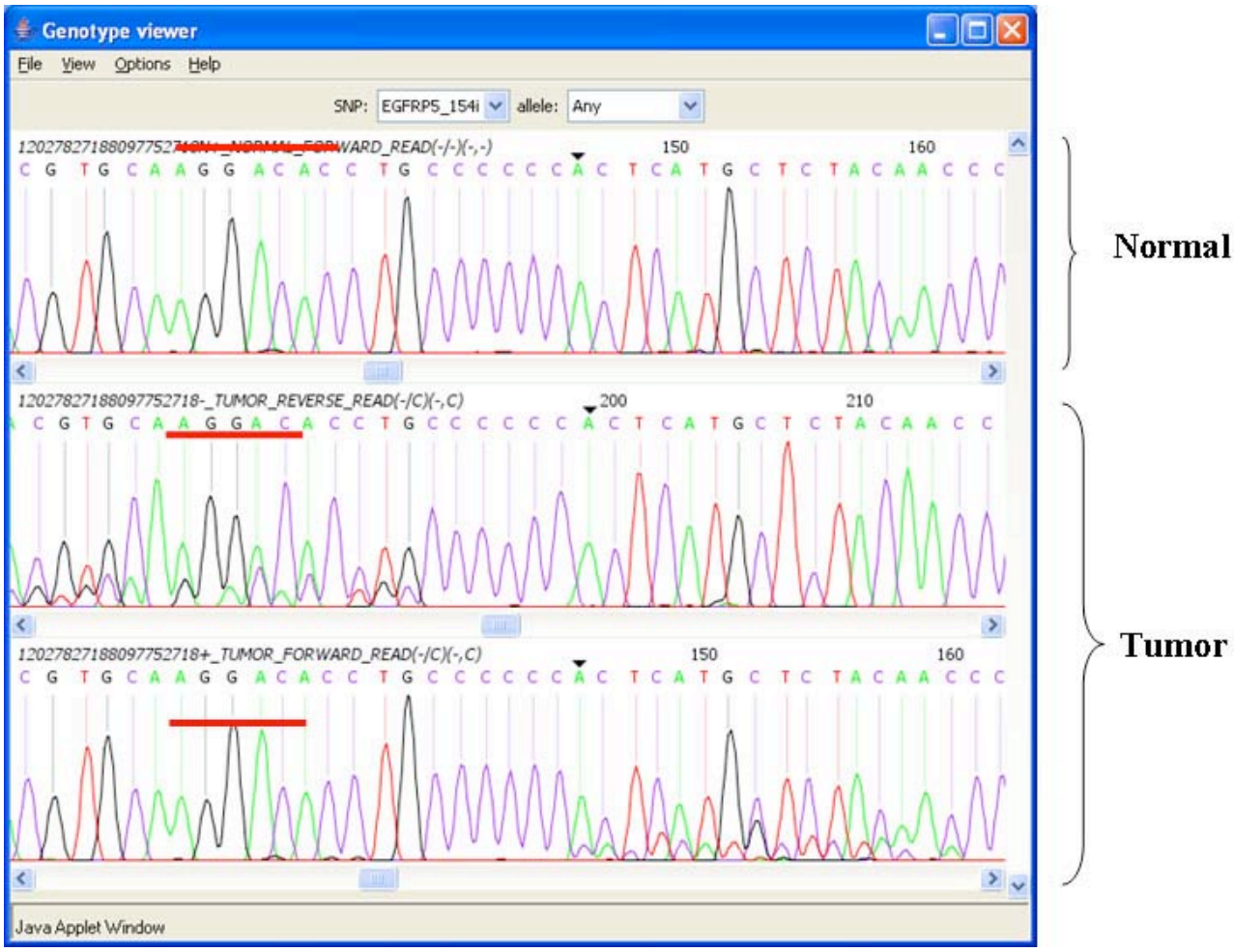


Unique users since Sept 1, 2010: 3408

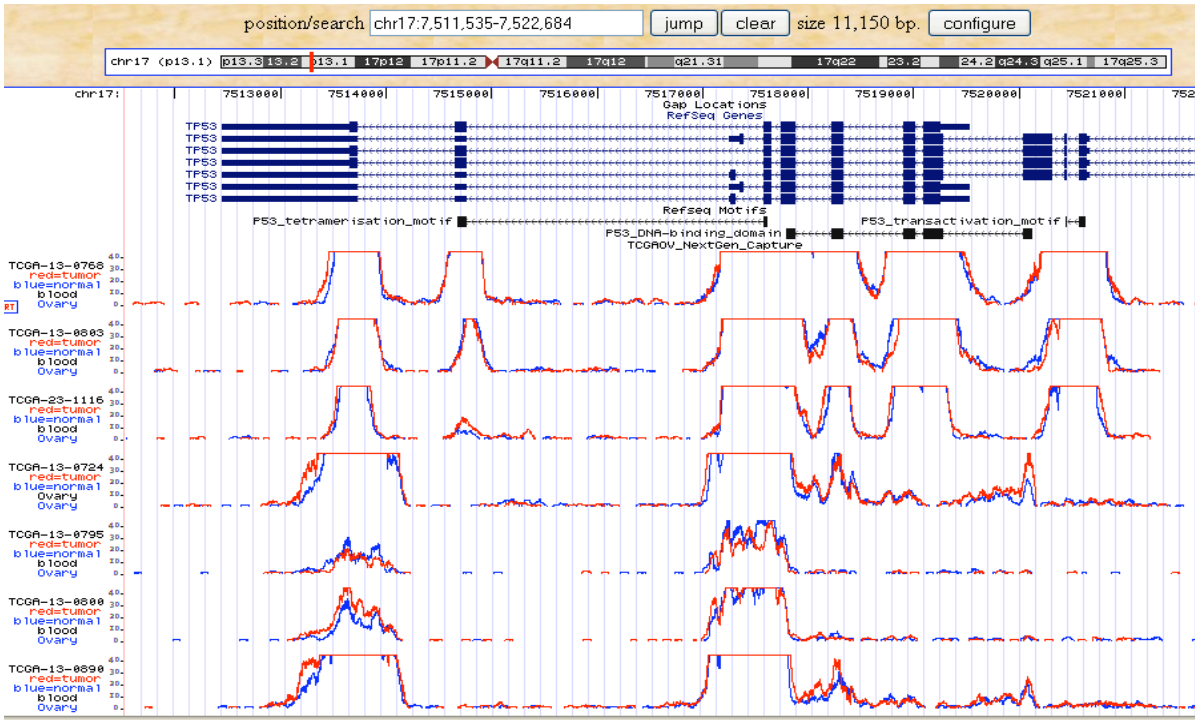
Comprehensive Genomic Summary



Putative Somatic Mutations can be Manually Reviewed
Eg: Frameshift Mutation in EGFR in Paired Tumor/Normal

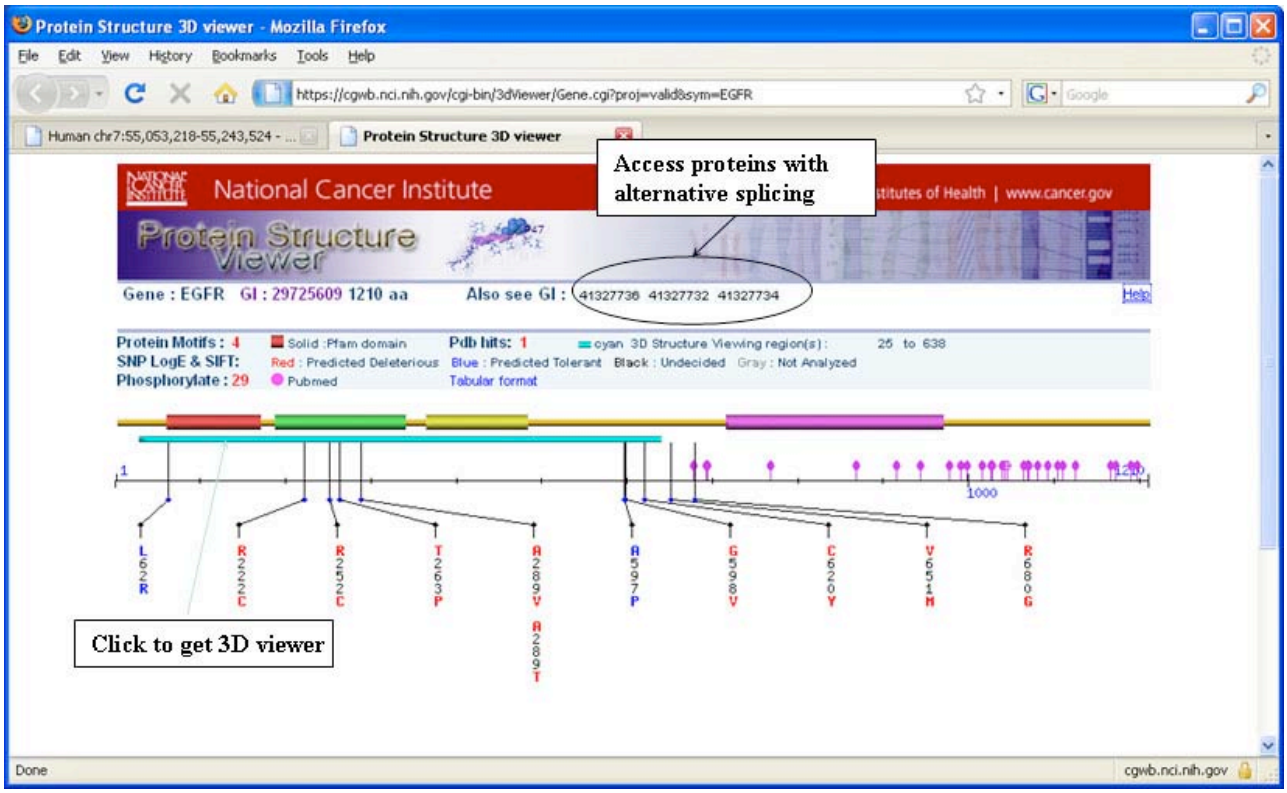

Genotype viewer
File View Options Help
SNP: EGFRP5_154i allele: Any
150
160
200
210
Java Applet Window
Normal
Tumor

Next-gen sequencing p53 analysis

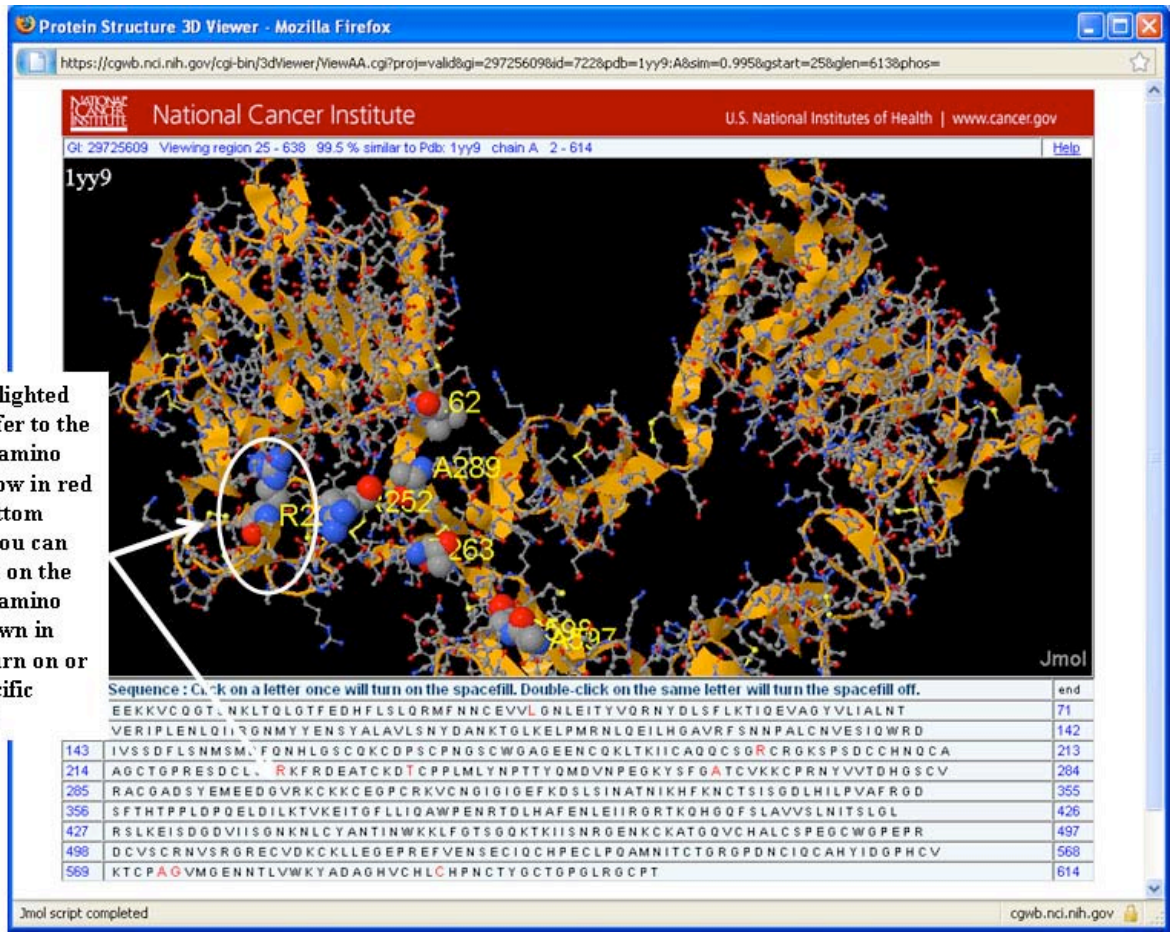


Showing coverage for its matching normal and coverage for a tumor sample
Sample coverage maximized to 40x, height represents coverage

Protein structure view of EGFR mutations

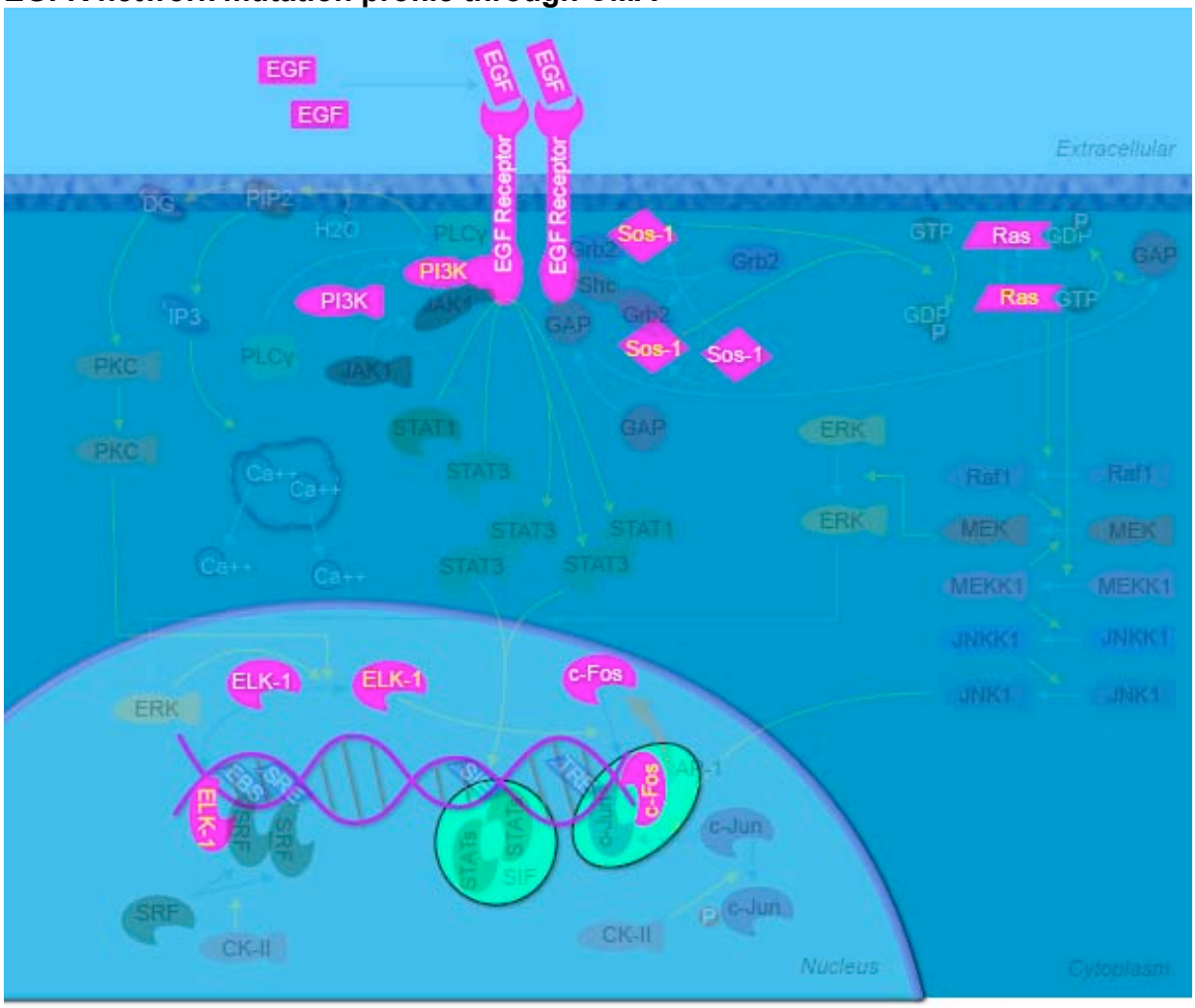


3D Structure Viewer

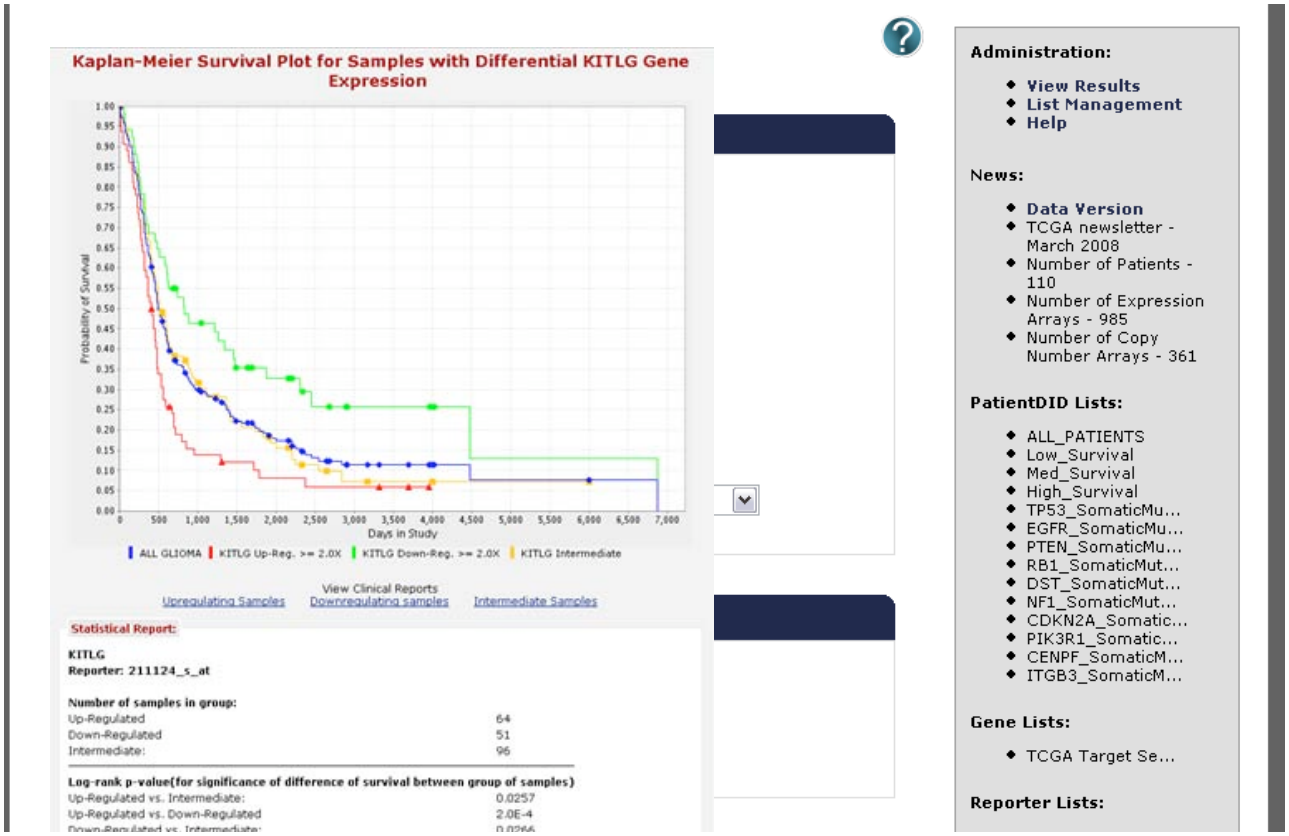


Big, highlighted atoms refer to the mutated amino acids (show in red in the bottom panel). You can also click on the mutated amino acid (shown in red) to turn on or off a specific mutation

EGFR network mutation profile through CMA

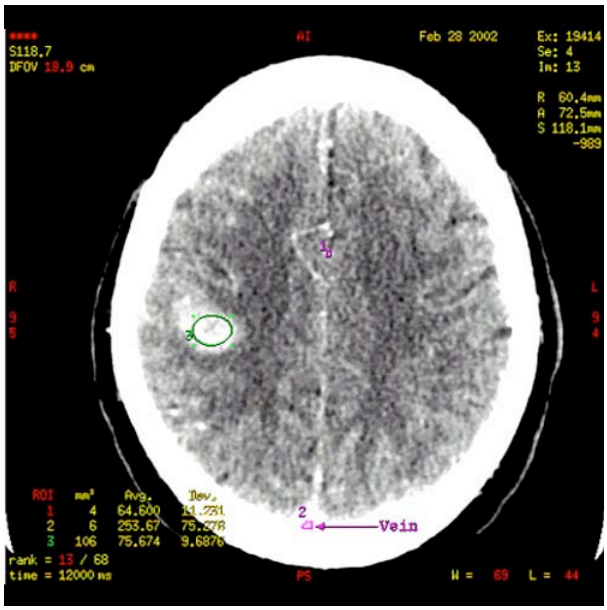


Gene expression analysis related to clinical outcome



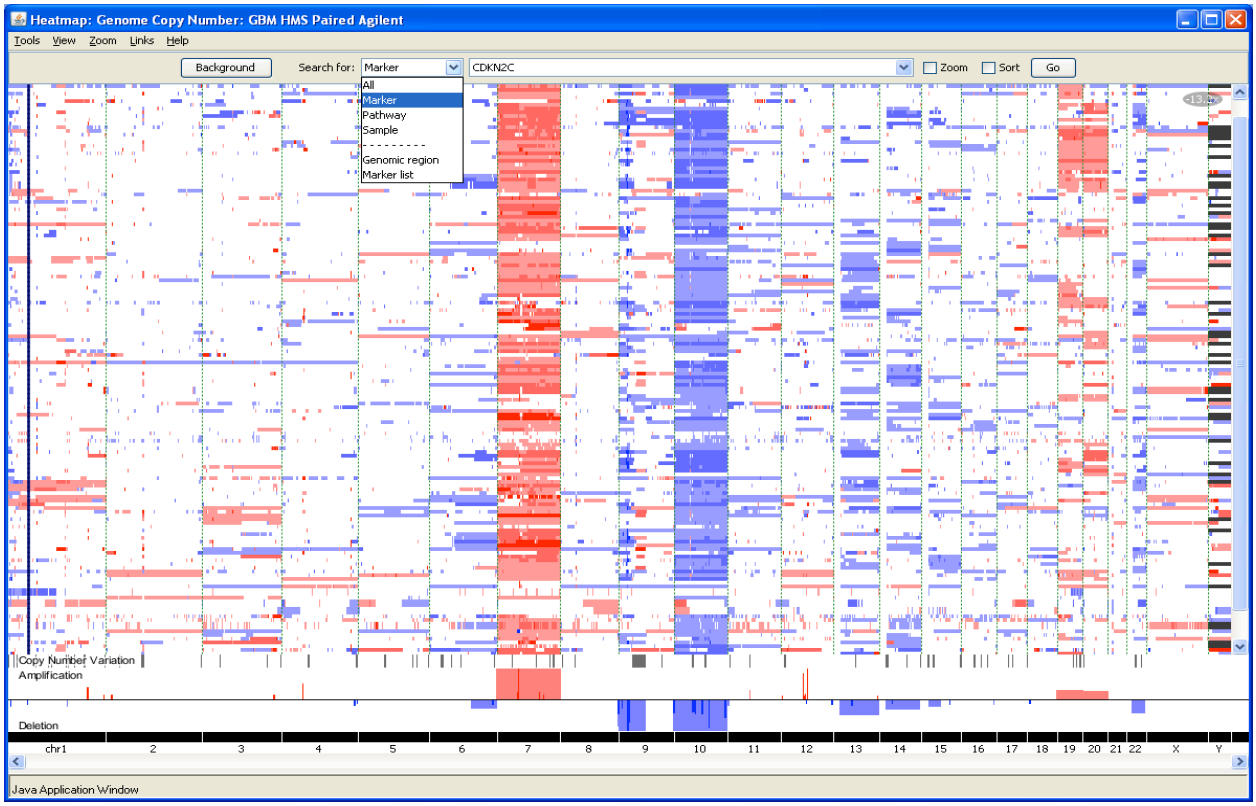
In silico research: hypothesis generation utilizing TCGA GBM resources

Glioblastoma Multiforme (GBM)

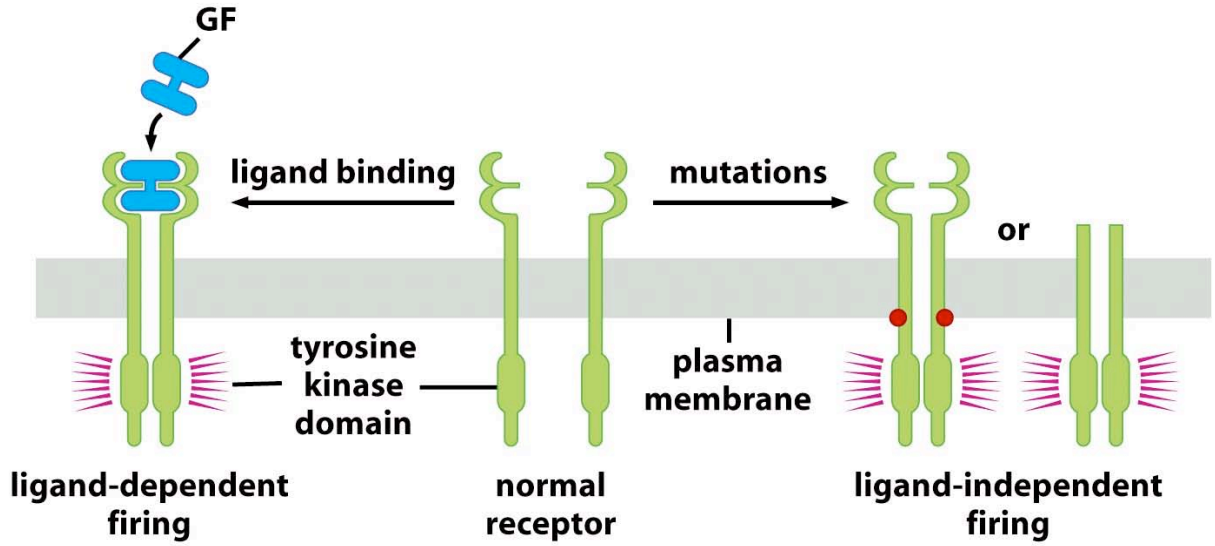


- GBM is the most common type of brain tumor. High grade gliomas are incurable and tumors expressing a mesenchymal phenotype are the most aggressive form

Chromosome 7 and EGFR seen as frequent targets of alteration in GBM



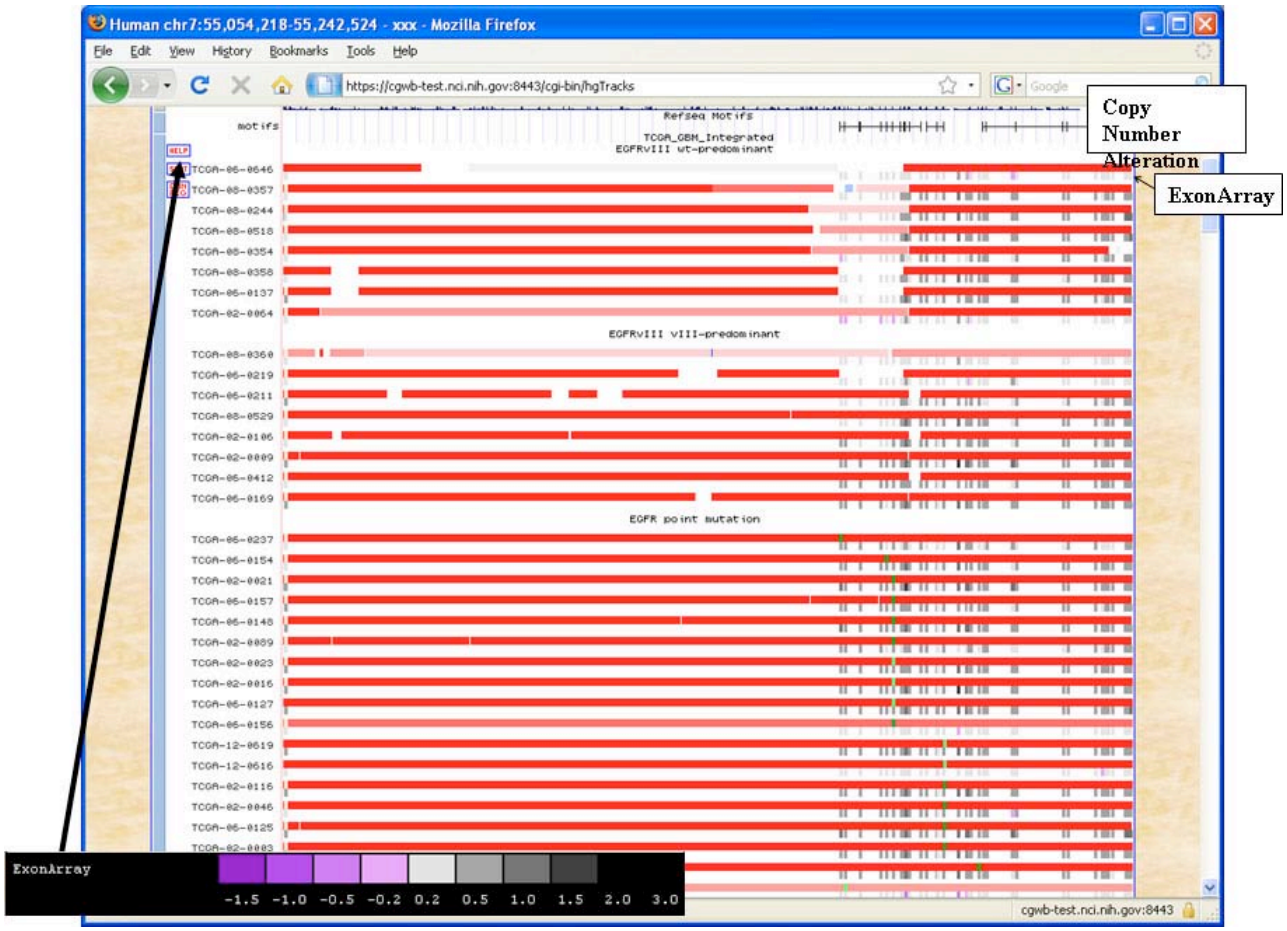
Constitutive activation of EGFR leads to abnormal growth



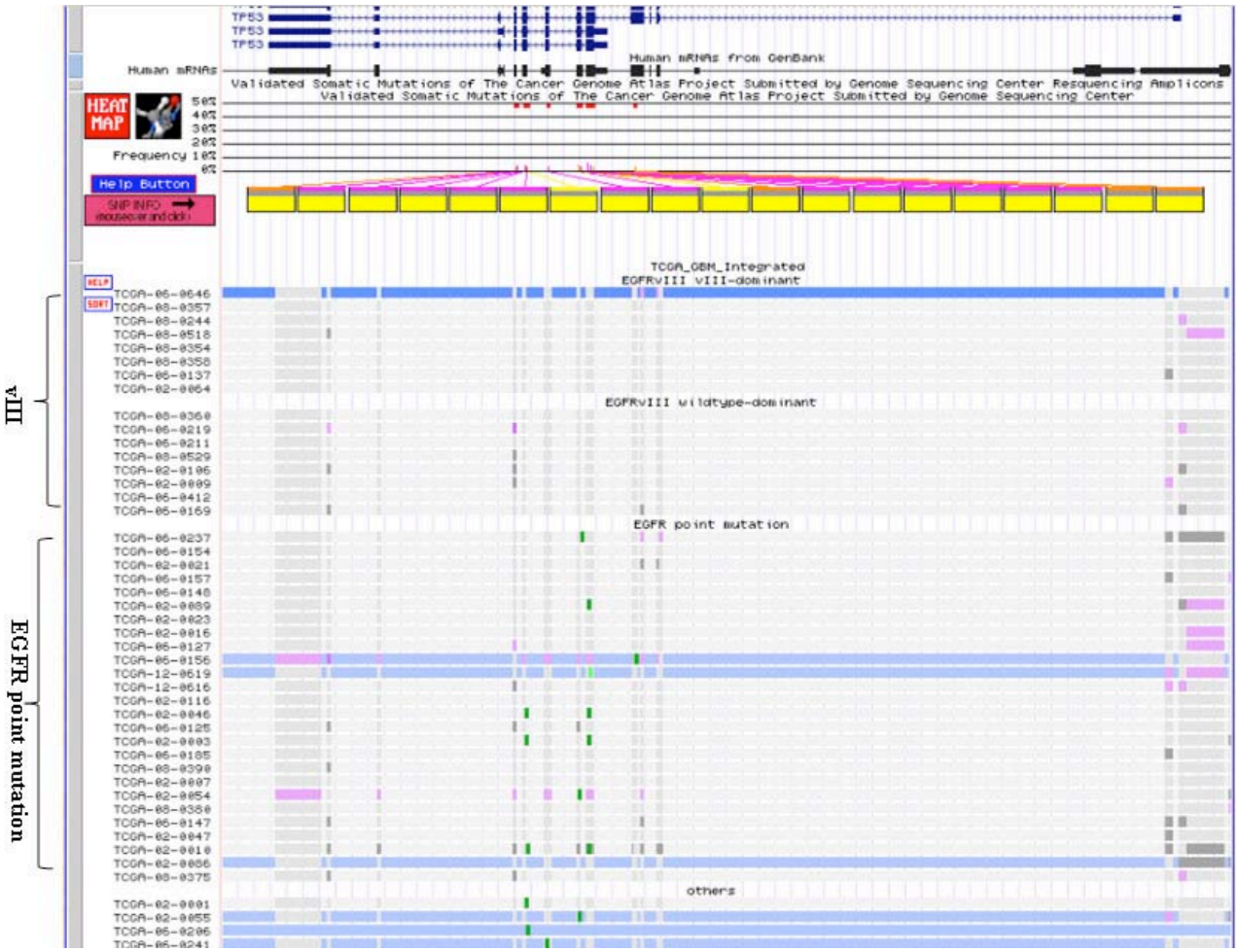
“EGFRvIII” mutation

Modified from The Biology of Cancer (© Garland Science 2007)

Add exon array to verify EGFRvIII expression correlates with CN
Top row: Copy Number + Somatic mutation + Methylation
Bottom row: Exon Array Expression



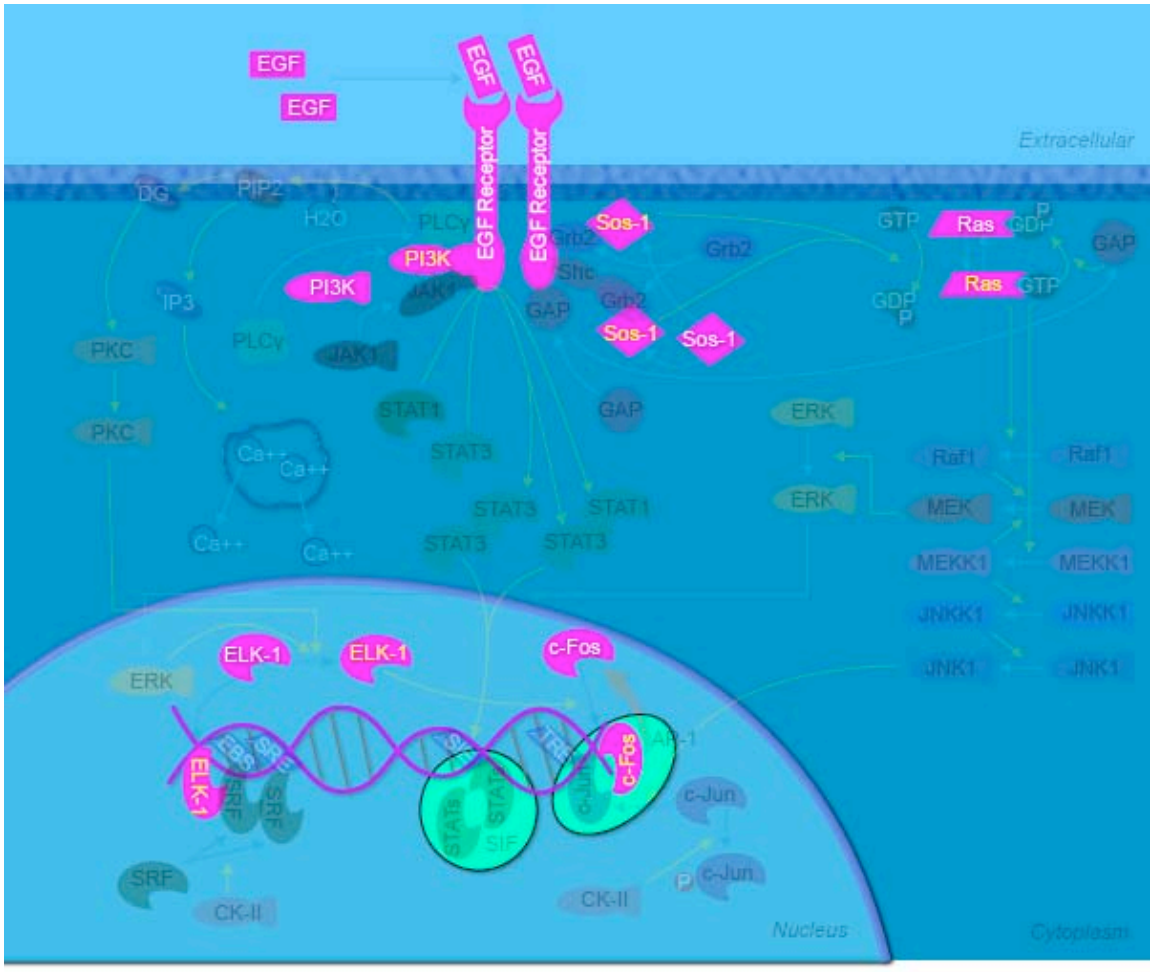
EGFR mutation subgroups viewed at the TP53 locus



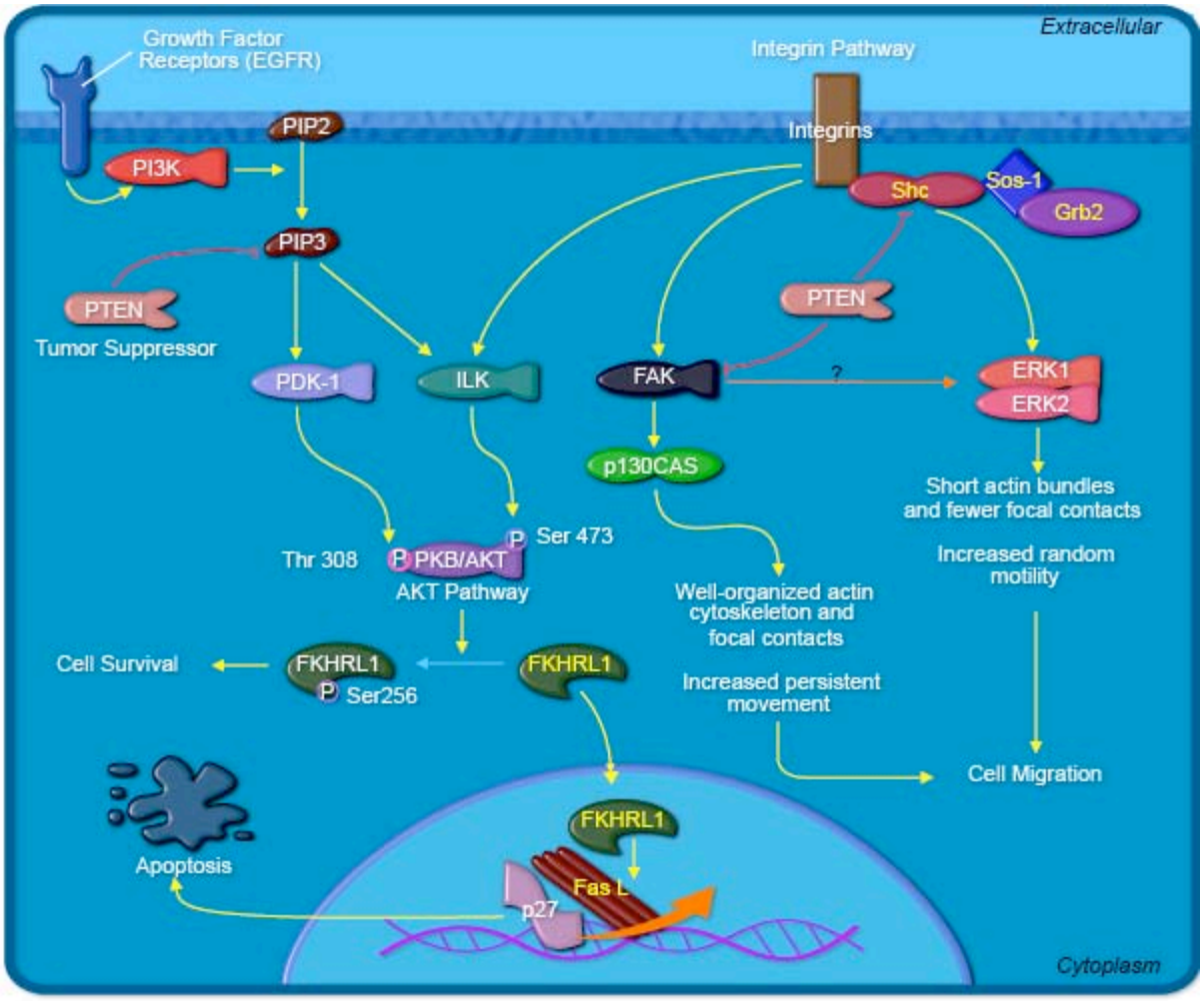
Mutations in EGFR vIII and TP53 may be anti-correlated

	EGFR amplification			No EGFR amplification		
	EGFR point mutation	EGFRvIII	No EGFR mutation	EGFR point mutation	EGFRvIII	No EGFR mutation
	18	12	37	7	0	79
TP53	5	0	4	3	0	35
Fraction:	28%	0%	11%	43%	N/A	44%

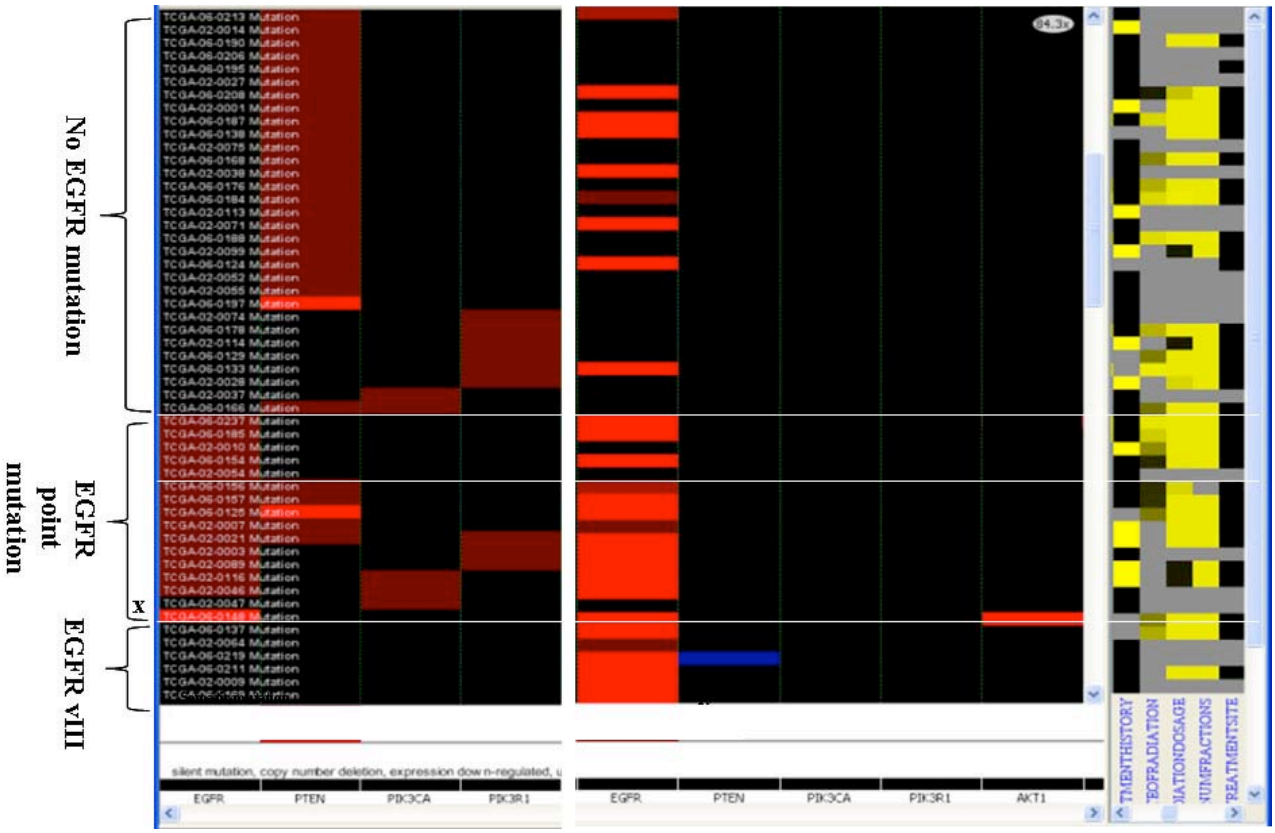
EGFR network mutation profile through CMA



Alterations in PI3K pathway through CMA



somatic mutations (left) and copy number (right) shows frequent co-occurrence of EGFR point mutations with other genes in PI-3K pathway but not the EGFR vIII mutations



In silico hypothesis

- No P53 mutations were found in amplified samples with EGFRvIII while significant levels of P53 mutation were found in amplified samples with EGFR point mutations. Suggests alternative molecular etiologies
- EGFR point mutations co-exist with additional mutations in other genes involved in PI-3K pathway while EGFRvIII rarely have additional mutations in PI-3K pathway. This suggests the possibility of oncogene addiction in EGFRvIII tumors but not in tumors with EGFR point mutations even though both types of mutations target EGFR extracellular domains.

Sample Use: Director's Challenge Lung Study

- Director's Challenge Consortium for the Molecular Classification of Lung Adenocarcinoma. Shedden K, Taylor JM, Enkemann SA, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site,blinded validation study. Nat Med. 2008 Aug;14(8):822-7. Epub 2008 Jul 20.
- "Here we report a large, training-testing, multi-site, blinded validation study to characterize the performance of several prognostic models based ongene expression for 442 lung adenocarcinomas. The hypotheses proposedexamined whether microarray measurements of gene expression eitheralone or combined with basic clinical covariates (stage, age, sex) could be used to predict overall survival in lung cancer subjects…This study alsoprovides the largest available set of microarray data with extensivepathological and clinical annotation for lung adenocarcinomas. "
- caBIG® Capabilities leveraged:
  - caArray
  - caIntegrator
  - caDSR

Sample Use: Molecular Dissection of Colon Cancer

- Sheffer M, Bacolod MD, et. al. Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer Proc Natl Acad Sci U S A. 2009;106(17):7131-6.
- "To identify and characterize chromosomal abnormalities in colorectal cancer, we performed a statistical analysis of 299 expression and 130 SNP arrays profiled at different stages of the disease, including normal tissue, adenoma, stages 1-4 adenocarcinoma, and metastasis."
- caBIG® Capabilities leveraged:
  - caArray
  - caIntegrator
  - caDSR

Patient selection for HER2 Tx required tissue screen and allowed only 1 of 4 women to participate

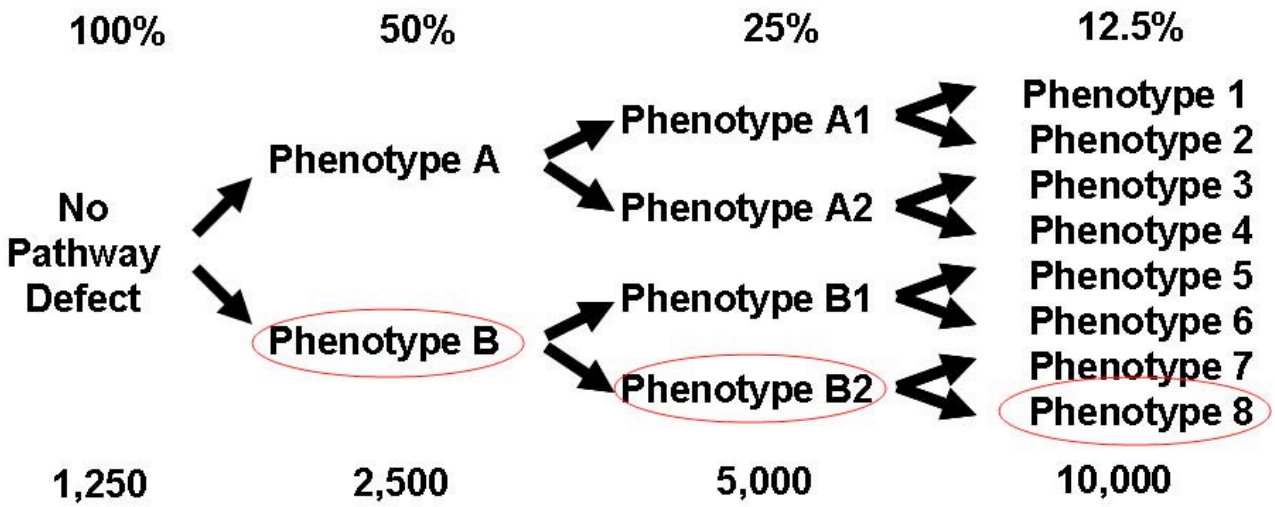
Calculated Sample Size And Study Duration		Hypothetical HER2+ Prevalence	Required "Screened" Population
1250	52 mos	100%	1250
		50%	2500
		25%	5000

- Need a obtain a suitable specimen, wait for test results. (Results were obtained in days to weeks)
- Need to screen many patients.

Courtesy H. Kim Lyerly, M.D., Director

Size of Population with Pathway to Inhibit*

Population fraction containing signature



Size of Population Needed To Screen

Courtesy H. Kim Lyerly, M.D., Director

The I-SPY TRIAL (Investigation of Serial studies to Predict Your Therapeutic Response with Imaging And moLecular analysis)

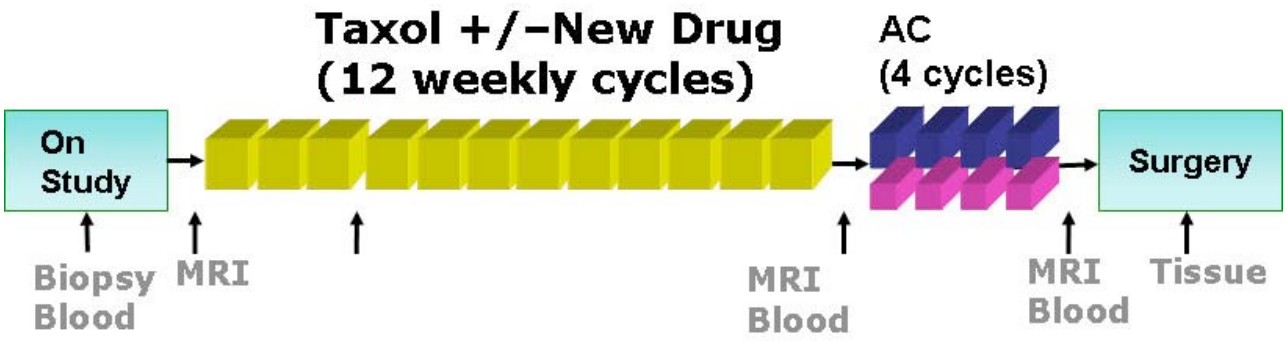
A national study to leverage biomarkers in predicting response to combinatorial therapy for women with Stage 3 breast cancer.

(PI Laura Esserman, UCSF)

Projected I-SPY 2 study sites

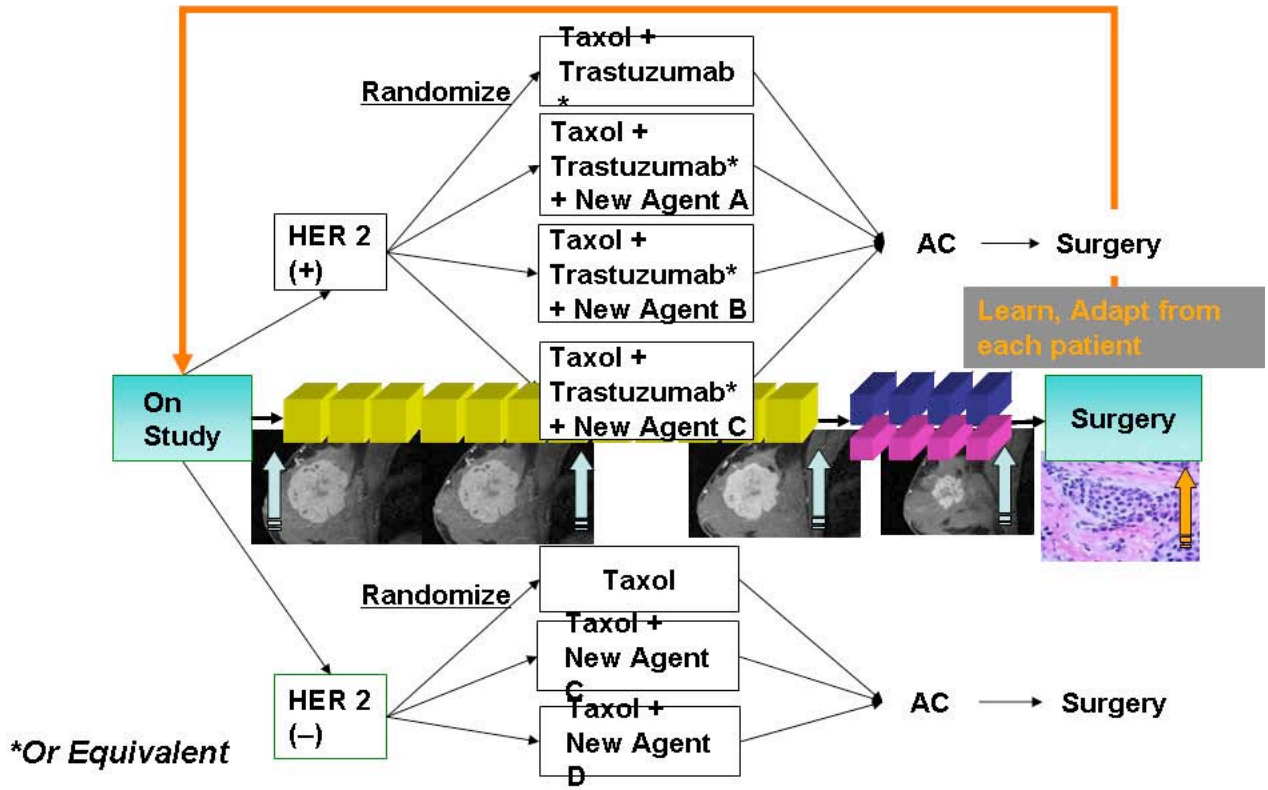
- Southwest Washington Medical Center
- Oregon Health & Science University
- Helen Diller Family Comprehensive Cancer Center
- University of California, San Diego Medical Center – Moores Cancer Center
- Mayo Clinic
- University of Colorado Cancer Center
- University of Kansas Medical Center
- Southwestern Medical Center
- University of Texas MD Anderson Cancer Center
- University of Chicago Cancer Research Center
- Emory
- University of Pennsylvania Abramson Cancer Center
- Georgetown Lombardi Comprehensive Cancer Center
- Inova Health System
- Emory Winship Cancer Institute
- University of Minnesota Medical Center

I-SPY Adaptive Trial Outline

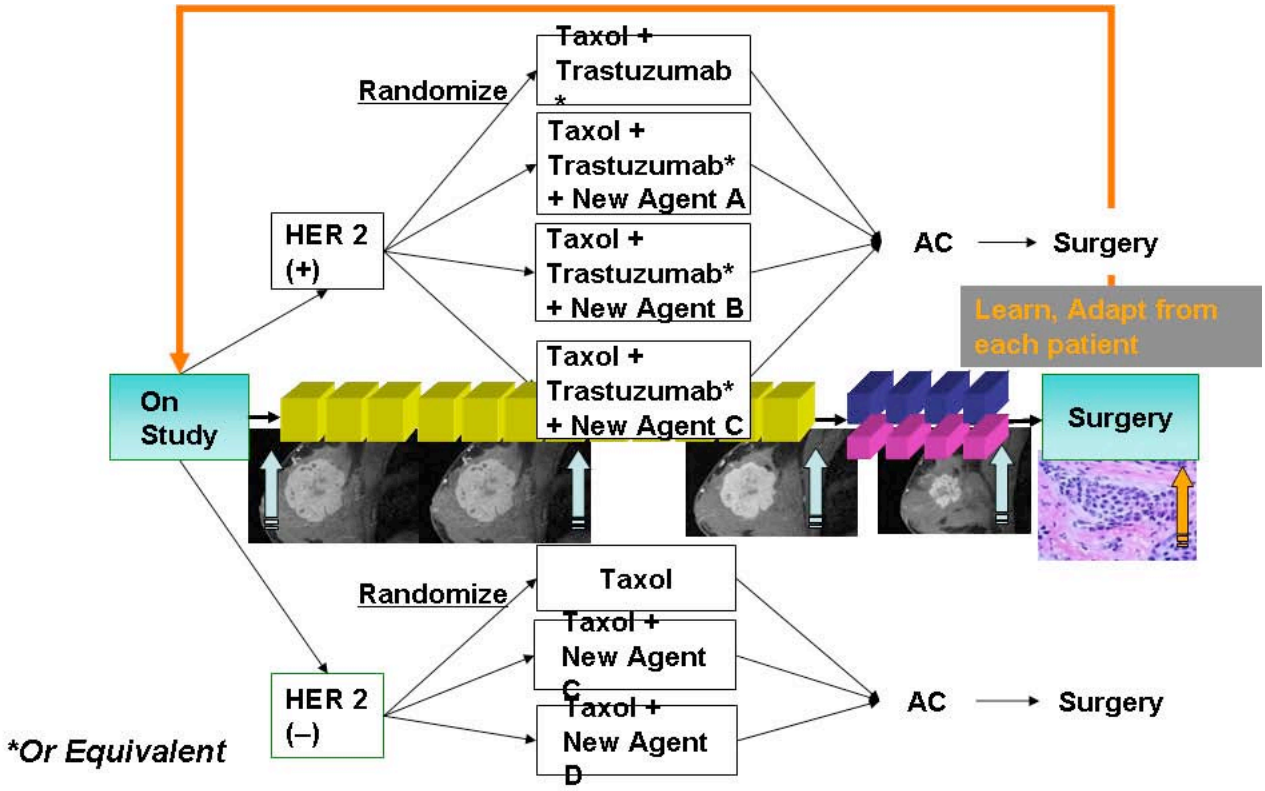


Accrual: Anticipate 800 patients over 3–4 years
Enroll ~20 patients per month
Participating Sites: 15–20 across US and Canada

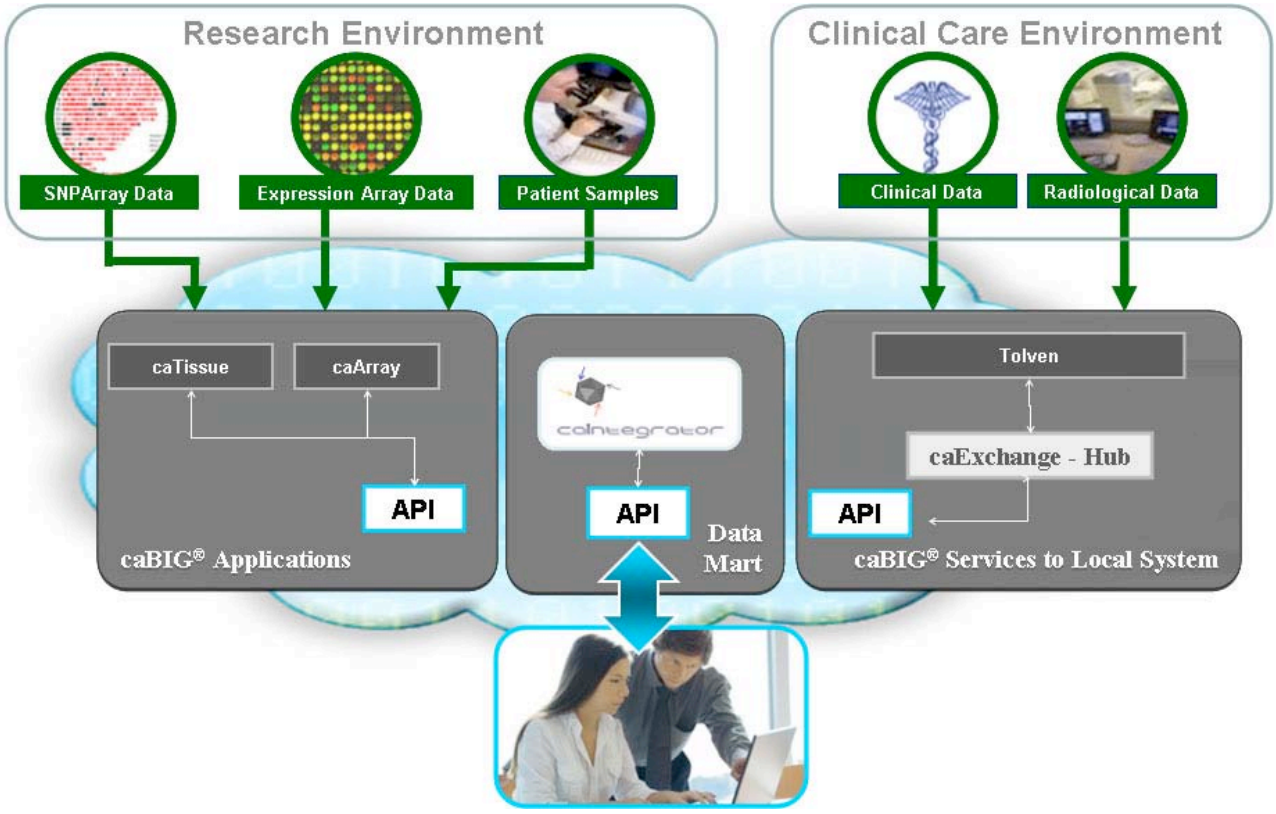
I-SPY Adaptive Trial:
Introduce several new agents for a given profile



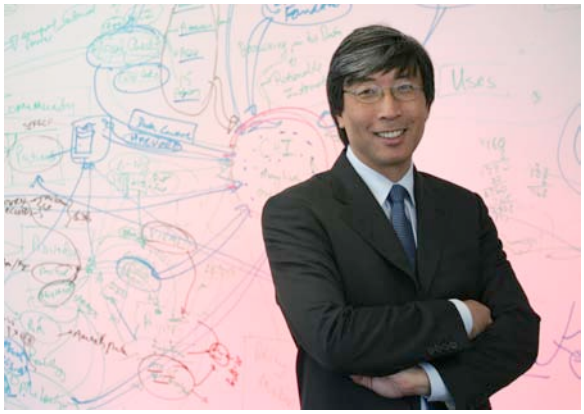
I-SPY Adaptive Trial:
Introduce several new agents for a given profile



I-SPY TRIAL IT Infrastructure



Our Distinguished Speaker



Patrick Soon-Shiong, M.D.
Executive Chairman
Abraxis Bioscience

The Vision for Personalized Medicine
and the Enabling Role of IT